Large Language Model guided Protocol Fuzzing

Year: 2024 | *Citations: 204* | *Authors: Ruijie Meng, Martin Mirchev, Marcel Böhme*

Abstract

—How to find security flaws in a protocol implementation without a machine-readable specification of the protocol? Facing the internet, protocol implementations are particularly security-critical software systems where inputs must adhere to a specific structure and order that is often informally specified in hundreds of pages in natural language (RFC). Without some machine-readable version of that protocol, it is difficult to automatically generate valid test inputs for its implementation that follow the required structure and order. It is possible to partially alleviate this challenge using mutational fuzzing on a set of recorded message sequences as seed inputs. However, the set of available seeds is often quite limited and will hardly cover the great diversity of protocol states and input structures. In this paper, we explore the opportunities of systematic interaction with pre-trained large language models (LLMs), which have ingested millions of pages of human-readable protocol specifications, to draw out machine-readable information about the protocol that can be used during protocol fuzzing. We use the knowledge of the LLMs about protocol message types for well-known protocols. We also checked the LLM’s capability in detecting “states” for stateful protocol implementations by generating sequences of messages and predicting response codes. Based on these observations, we have developed an LLM-guided protocol implementation fuzzing engine. Our protocol fuzzer C HAT AFL constructs grammars for each message type in a protocol, and then mutates messages or predicts the next messages in a message sequence via interactions with LLMs. Experiments on a wide range of real-world protocols from P RO F UZZ B ENCH show significant efficacy in state and code coverage. Our LLM-guided stateful fuzzer was compared with state-of-the-art fuzzers AFLN ET and NSF UZZ . C HAT AFL covers 47.60% and 42.69% more state transitions, 29.55% and 25.75% more